

Mutation screening and association study of the *TSSK4* gene in Chinese infertile men with impaired spermatogenesis

Running head: mutation and association study of *TSSK4*

Dan Su,^{*†} Wei Zhang,^{*†} Yuan Yang,^{*} Ying Deng,^{*} Yongxin Ma,^{*} Hongxia Song,^{*} Sizhong Zhang^{*}

^{*}Department of Medical Genetics, West China Hospital, Sichuan University, and Division of Human Morbid Genomics, State Key Laboratory of Biotherapy, P.R. China

[†]These authors contributed equally to this work

Correspondence to: Dr. Sizhong Zhang, Department of Medical Genetics, West China Hospital, Sichuan University, Renmin Nanlu, Section 3 #17 Chengdu, 610041, China

Email: szzhang@mcwcums.com

TEL: (86) 28-85164009

FAX: (86) 28-85164009

This work was supported by the National Natural Science Foundation of China (Grant No. 30470960) and the China Medical Board Foundation of New York.

ABSTRACT: Testis specific serine/threonine kinase (TSSK) family is a specific kinase group with exclusive or dominant expression in testis and involvement in spermatogenesis and male infertility. TSSK4 is a newly identified member of the TSSK family. In order to investigate the possible relations between variations including mutations and polymorphisms of the *TSSK4* gene and impaired spermatogenesis in human, mutation screening of this gene in 372 patients with azoospermia or severe oligospermia and 220 controls was performed. Total four novel single nucleotide changes including c.679G>A, c.987+108G>A, c.-155C>G and c.765C>A were discovered and the latter two variations were found only in patients. Bioinformatics analysis suggested that the allele A of c.765C>A could decrease the activity of pre-mRNA splicing of the *TSSK4*. The frequency of allele A of c.679G>A was significantly higher in controls than that in patients. On the contrary, allele A of c.987+108G>A was remarkably increased in patients compared to controls. Our investigation on the *TSSK4*, a potentially important testicular gene, in Chinese infertile and control men identified the association of some SNPs in this gene with male infertility.

Key words: male infertility, single nucleotide polymorphism (SNP), spermatogenesis

Introduction

A significant proportion of idiopathic human male infertility, usually accompanied by azoospermia or severe oligospermia, has been considered to have certain genetic defects (de Kretser 1997; Matzuk et al. 2002). Although hundreds of genes have been reported as possible candidates, very few genetic causes have been confirmed. The majority of the defects still need to be tested in patients with impaired spermatogenesis. Besides mutation, other genetic variations including single nucleotide polymorphisms (SNPs) have been increasingly noted, which may act as risk factors in modifying the susceptibility to male infertility (Krausz et al. 2007).

During spermatogenesis, protein phosphorylation is essential for signaling pathway and regulation of protein activity. Among protein kinases catalyzing phosphorylation, the testis specific serine/threonine kinase (TSSK) family is a particular one with exclusive or predominant expression in testis and functions in sperm differentiation, capacitation and fertilization (Bielke et al. 1994; Kueng et al. 1997; Zuercher et al. 2000; Hao et al. 2004; Chen et al. 2005; Xu et al. 2007). To date, five members of the TSSK family have been reported, including TSSK1, TSSK2, TSSK3, TSSK4 and TSSK6, and all of the human homologs have been cloned. The TSSK1 and TSSK2 are present in stage of late spermatid to sperm (Kueng et al. 1997; Hao et al. 2004; Xu et al. 2007). Also the TSSK3 exclusively expresses in human testis (Visconti et al. 2001), and in mature mice, *Tssk3* may function in the differentiated Leydig cells (Zuercher et al. 2000). The TSSK6, also named small serine/threonine kinase (SSTK), is found predominantly in the elongating spermatids and involved in postmeiotic chromatin remodeling (Hao et al. 2004; Spiridonov et al. 2005). Male mice with deficient *Tssk6* are infertile due to considerable sperm reduction, abnormal motility and morphology of spermatozoa (Spiridonov et al. 2005).

Recently, a new TSSK family member, TSSK4, initially named as TSSK5, has been identified. The *TSSK4* gene, mapped to 14q11.2 containing four exons, encodes a protein of 328 amino acids with specific expression in testis (Chen et al. 2005; Xu et al. 2007). Further study shows that TSSK4 phosphorylates the cAMP responsive element binding protein (CREB) at Ser-133 and hence facilitates the binding of CREB to the specific *cis* cAMP responsive element

(CRE), which is important for activating transcription of genes related to germ cell differentiation (Don et al. 2002; Chen et al. 2005). By stimulating the CRE/CREB pathway, *TSSK4* may be involved in spermatogenesis, probably in the early stages of spermatid differentiation (Don et al. 2002; Chen et al. 2005). Moreover, autosomal aberration analysis in patients with spermatogenic arrest indicated that band 14q11, at which *TSSK4* located, is highly linked with infertility suggesting that certain genes in this region may be associated with impaired spermatogenesis (Guo et al. 2002; Chen et al. 2005).

Therefore, to investigate the possible association between *TSSK4* and male infertility, we carried out a mutation screening of this gene in 372 infertile men with azoospermia or severe oligospermia and compared the results with those from 220 controls.

Materials and Methods

Subjects

The total 372 infertile patients, aged from 25 to 40 years, including 219 with azoospermia and 153 with severe oligospermia (sperm concentration $< 5 \times 10^6$ sperm/ml) were recruited from West China Hospital, Sichuan University. All of them underwent at least twice semen analyses according to World Health Organization guidelines (1999). Also their chromosomal abnormalities and microdeletions of azoospermia factor (AZF) region on Y chromosome were excluded by cytogenetic and corresponding sequence tagged site (STS) analysis, respectively (Simonim et al. 1999). The control group consisted of 220 fertile men with normospermia aged from 25 to 50. All participants in this study were Chinese Hans from Southwest of China and genetically unrelated. The study was approved by the Institutional Ethic Review Board of West China Hospital, Sichuan University, and informed consent was obtained from all subjects.

PCR Amplification

Genomic DNA was extracted from peripheral blood lymphocytes using standard phenol-chloroform procedures. Based on mRNA sequence (GenBank No.NM_174944.2) and

genomic sequence, four pairs of primers were designed using Primer Premier 5.0 to amplify all the four exons including intron/exon boundaries, 5' untranslated region (5' UTR) and 3'UTR (Table 1). PCRs were carried out in a total volume of 50 μ l containing 0.2 μ g genomic DNA, 10 pmol of each primer, 10 pmol of dNTP, and 2 units Taq polymerase (TaKaRa, Shiga Japan). The PCR profile was: pre-denaturation at 94°C for 5 min, 35 amplification cycles comprising denaturation at 94°C for 30 sec, annealing at a temperature between 53°C and 57°C for 30 sec, extension at 72°C for 1 min, with a final extra extension at 72°C for 10 min. Amplicons were resolved by 1.5% agarose gel electrophoresis to confirm the presence of specific PCR products.

Denaturing High-Performance Liquid Chromatography (DHPLC) Analysis and DNA Sequencing

Mutation screening of the *TSSK4* was carried out using denaturing high-performance liquid chromatography (DHPLC) on automated WAVE system (Transgenomic, Omaha, U.S.A.). The software WAVEMAKER 4.1 (Transgenomic) was used to determine optimal melting temperature for tested fragments. Prior to DHPLC, the PCR products were denatured at 94°C for 5 min and cooled at room temperature over 45 min. Then, 6 μ l products were injected into a high-throughput DNASep column and eluted with a linear acetonitrile gradient of 2% per minute at a flow rate of 0.9 ml/min. The elution profiles of heterozygous fragments were represented as aberrant shaped peaks while the homozygous fragments as single peaks. After DHPLC, heterozygous fragments were reamplified and purified with QIAquick PCR purification Kit (Qiagen, Hilden, Germany) and sent for direct sequencing in both directions on ABI3100 DNA Sequencer (Applied Biosystems, Foster City, USA).

Genotyping and Statistical Analysis

All participants were genotyped for variations identified including single nucleotide polymorphisms (SNPs) by PCR-restriction fragment length polymorphism (PCR-RFLP) analysis with corresponding restriction endonucleases (DdeI, BsaHI, Tsp45I, DpnII; NEB, Beverly, USA).

After endonuclease digestion, the products were electrophoresed on 3% agarose gel and observed with Gel Doc1000 system (Bio-Rad, Hercules, USA).

Hardy-Weinberg equilibrium was tested using HWE program. Differences in genotypic and allelic frequencies between infertile patients and controls were assessed by χ^2 test using software SPSS11.0.

Results

After screening of all four exons with their intron/exon boundaries, 5' UTR and 3'UTR of the *TSSK4* in 592 patients and controls, four single nucleotide changes were identified by DHPLC and DNA sequencing (Figure 1). According to the nomenclature recommendations of sequence variations (<http://www.hgvs.org/mutnomen>), they were named as c.-155C>G, c.679G>A, c.765C>A and c.987+108G>A. All of them had not been reported before. As shown in Figure 1, c.-155C>G and c.987+108G>A are located in 5' UTR and 3'UTR, respectively. In exon 3, the c.679G>A results in amino acid change of Val to Ile, whereas the c.765C>A does not alter the amino acid sequence of TSSK4. Both c.-155C>G and c.765C>A were detected only in five infertile men each, more precisely, in two azoospermic and three oligospermic patients for c.-155C>G and in five patients with azoospermia for c.765C>A. No patient possessed both of them simultaneously. In contrast to the two rare variations, c.679G>A and c.987+108G>A found in patients and controls should be considered as SNPs for that in both groups their minor allele frequencies were over 1%.

The genotypic distributions of the two SNPs were in Hardy-Weinberg equilibrium, and their genotype and allele frequencies were shown in Table 2. The two SNPs were found in both azoospermic and oligospermic patients without apparent differences in either genotype or allele frequencies between them (data not shown). As from the table, although no significant difference was observed among the three genotypes of c.679G>A, the frequency of allele A in controls was significantly higher than that in patients (3% versus 1.2%, $P=0.031$) with odds ratio (OR) of 0.400 (95% confidence interval [CI]: 0.170-0.944). At c.987+108G>A, only two genotype GG

and GA were identified in both groups. The frequencies of genotype GA and allele A decreased significantly in controls compared with patients (genotype GA: 15% versus 5.5%, $P < 0.001$; allele A: 7.5% versus 2.7%, $P = 0.001$), giving OR of 3.057 (95%CI: 1.600-5.842) and 2.89 (95%CI: 1.531-5.453), respectively.

Discussion

Although exact roles of the TSSKs are still unknown, a series of studies suggests that they are involved in the post-meiotic germ cell differentiation (Xu et al. 2007). Therefore it is reasonable to postulate that this kinase family may be associated with human male infertility and mutation screening of the TSSK genes in patients with impaired spermatogenesis may contribute to the understanding of their roles. In the present study, we screened the *TSSK4* variations and explored their possible association with spermatogenic impairment in human. Similar works on other four members of the *TSSK* family are in progress in our laboratory.

As a result, four novel variations of the *TSSK4* were identified. Although c.765C>A did not change the amino acid sequence, the result from bioinformatics analysis with ESEfinder 3.0 (Exonic Splicing Enhancer (ESE); <http://rulai.cshl.edu/cgi-bin/tools/ESE3/>; Cartegni et al. 2003) showed that the substitution of C by A at c.765C>A would result in disappearance of a ESE sequence GGTCACTT, which could be recognized and bound by the splicing factor SC35, an important member of serine/arginine-rich (SR) protein family that plays an essential role in pre-mRNA splicing (Cartegni et al. 2003; Sanford et al. 2005). Thus, compared to allele C, the allele A may decrease the activity of pre-mRNA splicing and reduce the expression of *TSSK4*. However, since c.765C>A is found in limited number of infertile men (5/372), whether it is related to impaired spermatogenesis still needs to be clarified by functional study in more patients. Another rare variation c.-155C>G should not be considered as a disease causing mutation for the time being, because its possible influence on gene function has not been studied.

At c.679G>A, significant difference of allele A frequency was observed between patients with oligo-/azoospermia and controls, suggesting this SNP may be associated with impaired

spermatogenesis. This SNP changes amino acid, but no effect on pre-mRNA splicing could be predicted, so it is not clear whether this polymorphism is a functional one or only a genetic marker being in linkage disequilibrium with other loci responsible for spermatogenic impairment. The statistical analysis also showed that the allele A of SNP c.987+108G>A was associated with increased risk for male infertility, since the frequencies of allele A (OR: 2.89) and genotype GA (OR: 3.057) were significantly higher in patients than those in controls. In addition, some researches have suggested that SNPs in the UTRs may impact mRNA stability, translation efficiency and gene expression (Miller et al. 2002; Nief et al. 2007; Kamiyama et al. 2007; Wang et al. 2007). Therefore, to further reveal the possible role of *TSSK4* in male infertility, more deep functional study of c.987+108G>A on gene expression is necessary.

In conclusion, as the first mutation analysis of the *TSSK4* in infertile men with azoospermia or severe oligospermia, our investigation provided some preliminary results and suggestions for in-depth studies of this gene on male infertility in the future.

References

- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21(2): 263-265.
- Bielke W, Blaschke RJ, Miescher GC, Zurcher G, Andres AC, Ziemiecki A. Characterization of a novel murine testis-specific serine/threonine kinase. 1994; *Gene* 139(2): 235-239.
- Bucko-Justyna M, Lipinski L, Burgering BM, Trzeciak L. Characterization of testis-specific serine-threonine kinase 3 and its activation by phosphoinositide-dependent kinase-1-dependent signalling. 2005; *FEBS J*. 272(24): 6310-6323.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*, 2003; 31: 3568-3571.
- Chen X, Lin G, Wei Y, Hexige S, Niu Y, Liu L, Yang C, Yu L. TSSK5, a novel member of the testis-specific serine/threonine kinase family, phosphorylates CREB at Ser-133, and stimulates the CRE/CREB responsive pathway. *Biochem Biophys Res Commun*. 2005; 333(3): 742-749.
- de Kretser DM. Male infertility. *The Lancet*. 1997; 349: 787-790.
- Don J, Stelzer G. The expanding family of CREB/CREM transcription factors that are involved with spermatogenesis. *Mol Cell Endocrinol*. 2002; 187(1-2): 115-124.
- Guo JH, Zhu PY, Huang YF, Yu L. Autosomal aberrations associated with testicular dysgenesis or spermatogenic arrest in Chinese patients. *Asian J Androl*. 2002; 4(1): 3-7.
- Hao Z, Jha KN, Kim YH, Vemuganti S, Westbrook VA, Chertihin O, Markgraf K, Flickinger CJ, Coppola M, Herr JC, Visconti PE. Expression analysis of the human testis-specific serine/threonine kinase (TSSK) homologues. A TSSK member is

- present in the equatorial segment of human sperm. *Mol Hum Reprod.* 2004; 10(6): 433-444.
- Kamiyama M, Kobayashi M, Araki S, Iida A, Tsunoda T, Kawai K, Imanishi M, Nomura M, Babazono T, Iwamoto Y, Kashiwagi A, Kaku K, Kawamori R, Ng DP, Hansen T, Gaede P, Pedersen O, Nakamura Y, Maeda S. Polymorphisms in the 3' UTR in the neurocalcin delta gene affect mRNA stability, and confer susceptibility to diabetic nephropathy. *Hum Genet.* 2007; 122(3-4): 397-407.
- Krausz C, Giachini C. Genetic Risk Factors in Male Infertility. *Arch Androl.* 2007; 53(3): 125-133.
- Kueng P, Nikolova Z, Djonov V, Hemphill A, Rohrbach V, Boehlen D, Zuercher G, Andres AC, Ziemiecki A. A novel family of serine/threonine kinases participating in spermiogenesis. *J Cell Biol.* 1997; 139(7): 1851-1859.
- Matzuk MM, Lamb DJ. Genetic dissection of mammalian fertility pathway. *Nat Med.* 2002; 8 (suppl 1): S41-49.
- Miller GM, Madras BK. Polymorphisms in the 3'-untranslated region of human and monkey dopamine transporter genes affect reporter gene expression. *Mol Psychiatry.* 2002; 7(1): 44-55.
- Nief N, Le Morvan V, Robert J. Involvement of gene polymorphisms of thymidylate synthase in gene expression, protein activity and anticancer drug cytotoxicity using the NCI-60 panel. *Eur J Cancer.* 2007; 43(5): 955-962.
- Sanford JR, Ellis J, Caceres JF. Multiple roles of arginine/serine-rich splicing factors in RNA processing. *Biochem Soc Trans.* 2005; 33: 443-446.
- Simonim M, Bakker E, Eurlings MCM, Matthijs G, Moro E, Muellers CR, Vogt PH. Laboratory guidelines for molecular diagnosis of Y-chromosomal microdeletions. *Int J Androl.* 1999; 22: 292-299.
- Spiridonov NA, Wong L, Zervas PM, Starost MF, Pack SD, Paweletz CP, Johnson GR. Identification and characterization of SSTK, a serine/threonine protein kinase essential for male fertility. *Mol Cell Biol.* 2005; 25(10): 4250-4261.
- Visconti PE, Hao Z, Purdon MA, Stein P, Balsara BR, Testa JR, Herr JC, Moss SB, Kopf GS. Cloning and chromosomal localization of a gene encoding a novel serine/threonine kinase belonging to the subfamily of testis-specific kinases. *Genomics.* 2001; 77(3): 163-170.
- Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nature Reviews Genetics.* 2007; 8(10): 749-761
- Xu B, Hao Z, Jha KN, Digilio L, Urekar C, Kim YH, Pulido S, Flickinger CJ, Herr JC. Validation of a testis specific serine/threonine kinase [TSSK] family and the substrate of TSSK1 & 2, TSKS, as contraceptive targets. *Soc Reprod Fertil Suppl.* 2007; 63: 87-101.
- Zuercher G, Rohrbach V, Andres AC, Ziemiecki A. A novel member of the testis specific serine kinase family, tssk-3, expressed in the Leydig cells of sexually mature mice. *Mech Dev.* 2000; 93(1-2): 175-177.

Table 1 *The PCR amplified region, primer sequence, annealing temperature, product size, and the DHPLC melting temperature for mutation screening of the TSSK4 gene*

Amplified Region	Primer Sequence (5'-3')	Annealing Temperature (°C)	Product Size (bp)	Melting Temperature (°C)
5'UTR+	F: CTTCTCTTTCTTTCCCCCTC	55.8	450	59.2
Exon 1	R: GCTCCAGTTCCTCTTCCAG			
Exon 2	F: GGGGTCAGGTAATGAAAGT R: GAAAAAAGGACAGGAAATGG	56.9	304	60.1
Exon 3	F: AAATAAGCCCTCTCCCCAG R: AGCCCTCCTTAGGTAGCCA	56.5	437	60
Exon 4+	F: GGAGAAAAACTCAGGCAAG	53.3	427	56.5
3'UTR	R: CAAGATTATTGGTTACAGGC			

Table 2 Genotype and allele frequencies of the four novel variations in the TSSK4 gene of infertile patients and controls

SNP (Location)			Controls (n=220)	Patients (n=372)	<i>P</i>
Restriction Endonucleases					
c.-155C>G (5'UTR) DdeI	Genotype	CC	220 (1)	367 (98.7)	
		CG	0 (0)	5 (1.3)	<i>NS</i>
		GG	0 (0)	0 (0)	
		CG+GG	0 (0)	5 (1.3)	<i>NS</i>
	Allele	C	440 (1)	739 (99.3)	
Frequency	G	0 (0)	5 (0.7)	<i>NS</i>	
c.679G>A (Exon3) BsaHI	Genotype	GG	208 (94.5)	363 (97.6)	
		GA	11 (5)	9 (2.4)	<i>NS</i>
		AA	1 (0.5)	0 (0)	<i>NS</i>
		GA+AA	12 (5.5)	9 (2.4)	<i>NS</i>
	Allele	G	427 (97.0)	735 (98.8)	
Frequency	A	13 (3.0)	9 (1.2)	0.031	
c.765C>A (Exon3) Tsp45I	Genotype	CC	220 (1)	367 (98.7)	
		CA	0 (0)	5 (1.3)	<i>NS</i>
		AA	0 (0)	0 (0)	
		CA+AA	0 (0)	5 (1.3)	<i>NS</i>
	Allele	C	440 (1)	739 (99.3)	
Frequency	A	0 (0)	5 (0.7)	<i>NS</i>	
c.987+108G>A (3'UTR) DpnII	Genotype	GG	208 (94.5)	316 (85.0)	
		GA	12 (5.5)	56 (15.0)	<0.001
		AA	0 (0)	0 (0)	
		GA+AA	12 (5.5)	56 (15.0)	<0.001
	Allele	G	428 (97.3)	688 (92.5)	
Frequency	A	12 (2.7)	56 (7.5)	0.001	

Data of genotype and allele frequency are presented as "n (%)". NS, not significantly

Figure legend

Figure 1. The single nucleotide variations found in the *TSSK4* gene. (A) Genomic structure of the *TSSK4* and the locations of the four variations. (B) DNA sequences of the four variations: the mutant sequences (upper panel) and the wild-type sequences (WT; lower panel). Arrows indicate the positions of the variations and codons are underlined by dark horizontal lines. The nomenclature of variations follows to the recommendations in web site <http://www.hgvs.org/mutnomen/>.

FIGURE 1, Su et al.

